

Math 151 10/1.

Conditional Probabilities and Random Variables

$$P_{X|A}(x) = P(X=x | A) = \frac{P((X=x) \cap A)}{P(A)}.$$

Since $\sum_x P((X=x) \cap A) = 1$.

(sum over all possible values of X).

$$\sum_x P_{X|A}(x) = 1.$$

e.g. In Bernoulli trials w/ prob p of success,
given that the 1st success occurs
on or before the n^{th} trial, what is
the p.m.f. of the # of trials until the
(1st success)

s.ln: Let $X = \# \text{ trials until 1st success}$.

and A be the event

$A = \{1\text{st success on or before } n^{\text{th}} \text{ trial}\}$

$$\begin{aligned} P(X=k | A) &= \frac{P((X=k) \cap A)}{P(A)} = \frac{P(X=k)}{P(A)} \\ &= 0 \text{ if } k > n \\ &\quad \text{else.} \end{aligned}$$

$$P(X=k) = (1-p)^{k-1} p.$$

$$P(A) = \sum_{k=1}^n (1-p)^{k-1} p.$$

$$\therefore P(X=k|A) \equiv P_{X|A}(k) = \frac{(1-p)^{k-1}}{\sum_{j=1}^n (1-p)^{j-1}}$$

$$= \frac{(1-p)^{k-1}}{\sum_{j=0}^{n-1} (1-p)^j} = \frac{(1-p)^{k-1}}{\frac{1-(1-p)^n}{1-(1-p)}} = \frac{p(1-p)^{k-1}}{1-(1-p)^n}$$

Conditioning one random variable on another

$$P_{X|Y}(x,y) \equiv P(X=x|Y=y)$$

$$= \frac{P((X=x) \cap (Y=y))}{P(Y=y)} = \frac{P_{X,Y}(x,y)}{P_Y(y)}$$

e.g. Problem 2.31.

4 independent rolls of a 6 sided die

$$X = \# \text{ of } 1's$$

$$Y = \# \text{ of } 2's.$$

Find the joint P.M.F of (X, Y) .

$$\begin{aligned}
 P(X=k, Y=j) &= P(X=k | Y=j) P(Y=j) \\
 &= P(X=k | Y=j) \binom{4}{j} \left(\frac{1}{6}\right)^j \left(\frac{5}{6}\right)^{4-j} \\
 &= \binom{4-j}{k} \left(\frac{1}{5}\right)^k \left(\frac{4}{5}\right)^{4-j-k} \cdot \binom{4}{j} \left(\frac{1}{6}\right)^j \left(\frac{5}{6}\right)^{4-j}
 \end{aligned}$$

note that $0 \leq k \leq 4-j$ here.
 $0 \leq j \leq 4$.

and $P(X=k, Y=j) = 0$ else.

Law of total probability for P.M.F's

$$P_X(x) = \sum_y P_{X,Y}(x,y) = \sum_y P_Y(y) P_{X|Y}(x|y).$$

Conditional expectations

Given an event A with $P(A) > 0$,

$$E[X|A] = \sum x P_{X|A}(x)$$

$$\text{and } E[g(x)|A] = \sum g(x) P_{X|A}(x).$$

(same calculation as before since
 $P_{X|A}(x)$ is a p.m.f.)

if A_1, \dots, A_n are disjoint w/ $P(A_i) > 0$
 $i = 1, \dots, n$ and $\bigcup_{i=1}^n A_i = \Omega$ (the sample space).

then

$$\begin{aligned}
 E(X) &= \sum_x x P_X(x) \\
 &= \sum_x \sum_{i=1}^n P((x=x) \cap A_i) \\
 &= \sum_x \sum_{i=1}^n P(x=x | A_i) P(A_i) \\
 &= \sum_{i=1}^n P(A_i) \sum_x x P(x=x | A_i) \\
 &= \sum_{i=1}^n P(A_i) E[X | A_i].
 \end{aligned}$$

$E(X) = \sum_{i=1}^n P(A_i) E(X | A_i)$

if the A_i are ($r=y$) for some r
we get.

$$E(X) = \sum_{i=1}^n P_r(y) E(X | r=y)$$

e.g. an alternative computation for mean
and variance of a geometric R.V.

X geometric w/ prob. P .

$$E(X) = E(X|X=1) \cdot P(X=1)$$

$$+ E(X|X>1) \cdot P(X>1)$$

$$= p + E(X|X>1)(1-p).$$

$$\text{but } E(X|X>1) = 1 + E(X).$$

(throw away the 1st try).

$$\text{so } E(X) = p + (1+E(x))(1-p).$$

$$= p + (1-p) + E(x) - pE(x).$$

$$\Rightarrow E(x) = \frac{1}{p}.$$

Similarly

$$E(X^2) = E(X^2|X=1)P_X(1) + E(X^2|X>1)P(X>1)$$

$$= p + E(X^2|X>1)(1-p).$$

that

$$\begin{aligned} E(X^2 | X > 1) &= E((X+1)^2) \\ &= E(X^2) + 2E(X) + 1. \end{aligned}$$

so

$$\begin{aligned} E(X^2) &= p + (E(X^2) + 2(\frac{1}{p}) + 1)(1-p) \\ &= p + E(X^2) - pE(X^2) + \frac{2(1-p)}{p} + 1-p. \\ \Rightarrow E(X^2) &= 1 + 2\frac{(1-p)}{p^2} + \frac{1-p}{p}. \\ &= \frac{2}{p^2} - \frac{2}{p} + \frac{1}{p} = \frac{2}{p^2} - \frac{1}{p} \end{aligned}$$

$$\text{and } \text{Var}(X) = E(X^2) - E(X)^2$$

$$= \frac{2}{p^2} - \frac{1}{p} - (\frac{1}{p})^2 = \frac{1}{p^2} - \frac{1}{p} = \frac{-p}{p^2}.$$

Independence of Random Variables

X a r.v. on Ω

A an event in Ω .

Def: X is independent of A iff
for each x in the range of X .

$$P((X=x) \cap A) = P(X=x) P(A).$$

if $P(A) > 0$, this is the

name as.

$$P(X=x | A) P(A) = P_X(x) P(A)$$

$$\therefore P(X=x | A) = P_X(x).$$

Def: Two random variables X, Y are independent iff

$$P((X=x) \cap (Y=y)) = P(X=x) P(Y=y)$$

$\forall x, y$ in ranges of X, Y .

iff $P_{x,y}(x,y) = P_x(x) P_y(y)$.

Similarly,

X, Y are conditionally independent given the event A iff

$$P(X=x, Y=y | A) = P(X=x | A) P(Y=y | A)$$
$$\forall (x, y).$$

If X, Y are independent r.v's
on the same sample space Ω
then

$$E(XY) = \sum_{(x,y)} xy P_{x,y}(x,y) = \sum_x \sum_y xy P_x(x) P_y(y)$$

$$= \sum_x x P_x(x) \sum_y y P_y(y) = E(X) E(Y)$$

if X, Y are independent then so are
 $g(X)$ and $h(Y)$ for funcs g, h .

$$\begin{aligned}
 & P(g(X)=y, h(Y)=z) \\
 &= P(X \in \{x : g(x)=y\}, Y \in \{y : h(y)=z\}) \\
 &= P\left(\bigcup_{x:g(x)=y} (X=x) \cap \bigcup_{w:h(w)=z} (Y=w)\right) \\
 &= P\left(\bigcup_{\{(x,y) : g(x)=y, h(y)=z\}} (X=x) \cap (Y=y)\right) \\
 &= \sum_{\{(x,y) : g(x)=y, h(y)=z\}} P((X=x) \cap (Y=y)) \\
 &= \sum_{(\text{same})} P_X(x) P_Y(y) = \sum_{x:g(x)=y} P_X(x) \sum_{y:h(y)=z} P_Y(y) \\
 &= P(g(X)=y) P(h(Y)=z)
 \end{aligned}$$

If X, Y are independent then

$$\text{Var}(X+Y) = E((X+Y)^2) - [E(X+Y)]^2$$

$$= E(X^2 + 2XY + Y^2) - (E(X)^2 + 2E(X)E(Y) + E(Y)^2)$$

$$= E(X^2) + 2E(XY) + E(Y^2) - (E(X)^2 + 2E(X)E(Y) + E(Y)^2)$$

$$= E(X^2) - E(X)^2 + E(Y^2) - E(Y)^2$$

$$+ 2(E(XY) - E(X)E(Y)).$$

$\underbrace{\quad}_{=0 \text{ by independence.}}$

Def'n $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$

is a measure of how "dependent"

X, Y are.

Note that $\text{Cov}(X, X) = \text{Var}(X)$.

~~By induction if X_1, \dots, X_n are independent~~

Caution $\text{Cov}(X, Y) = 0$ does not imply that X, Y are independent

A collection of events

A_1, \dots, A_n are said to be independent

if

$$P\left(\bigcap_{i \in S} A_i\right) = \prod_{i \in S} P(A_i)$$

for every subset S of $\{1, 2, \dots, n\}$

Random variables X_1, \dots, X_n are independent if for each

choice of $\{x_1, \dots, x_n\}$ where $x_i \in \text{Range}(X_i)$

$$\{X_1 = x_1\}, \dots, \{X_n = x_n\}$$

are independent events.

equivalently

$$P_{X_1, \dots, X_n}(x_1, \dots, x_n) = P_{X_1}(x_1) \dots P_{X_n}(x_n)$$

✓ choices (x_1, \dots, x_n) .

Also, and for any functions $f(x_1), \dots, f_n(x_n)$

are ind. $f(X_1, x_2, \dots, x_n), g(X_{n+1}, \dots, x_n)$ are ind etc

By induction, if X_1, \dots, X_n are independent then

$$\text{Var}(X_1 + \dots + X_n) = \text{Var}(X_1) + \dots + \text{Var}(X_n).$$

Here's another way to see it.

$$(X_1 + \dots + X_n)^2 = \sum_{i=1}^n \sum_{j=1}^n X_i X_j$$

$$\text{So } E[(X_1 + \dots + X_n)^2] = \sum_{i=1}^n \sum_{j=1}^n E(X_i X_j)$$

$$= \sum_{i=1}^n E(X_i^2) + \underbrace{\sum_{i \neq j} E(X_i) E(X_j)}_{\text{by independent}}$$

$$\text{Also } (E(X_1 + \dots + X_n))^2 = \sum_{i=1}^n \sum_{j=1}^n E(X_i) E(X_j) \quad (\text{linearity}).$$

$$= \sum_{i=1}^n E(X_i)^2 + \sum_{i \neq j} E(X_i) E(X_j).$$

$$\text{So } E[(X_1 + \dots + X_n)^2] - (E(X_1 + \dots + X_n))^2$$

$$= \sum_{i=1}^n E(X_i^2) - \sum_{i=1}^n E(X_i)^2$$

$$= \sum_{i=1}^n \text{Var}(X_i)$$

e.g. binomial
= sum ind. Bernoulli
e.g. negative binomial
= sum ind. geometric